Journal of Statistical Planning and Inference 140 (2010) 2632–2652

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Kernel estimators for the second order parameter in extreme value statistics

Yuri Goegebeur [a,*,1], Jan Beirlant [b], Tertius de Wet [c,2]

[a] Department of Mathematics and Computer Science, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark
[b] Department of Mathematics and Leuven Statistics Research Centre, K.U.Leuven, Celestijnenlaan 200B, 3001 Heverlee, Belgium
[c] Department of Statistics and Actuarial Science, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa

ARTICLE INFO

Article history:
Received 15 September 2009
Received in revised form
12 February 2010
Accepted 8 March 2010
Available online 19 March 2010

Keywords:
Extreme value statistics
Pareto-type model
Second order parameter
Kernel statistic

ABSTRACT

We develop and study in the framework of Pareto-type distributions a general class of kernel estimators for the second order parameter ρ, a parameter related to the rate of convergence of a sequence of linearly normalized maximum values towards its limit. Inspired by the kernel goodness-of-fit statistics introduced in Goegebeur et al. (2008), for which the mean of the normal limiting distribution is a function of ρ, we construct estimators for ρ using ratios of ratios of differences of such goodness-of-fit statistics, involving different kernel functions as well as power transformations. The consistency of this class of ρ estimators is established under some mild regularity conditions on the kernel function, a second order condition on the tail function $1-F$ of the underlying model, and for suitably chosen intermediate order statistics. Asymptotic normality is achieved under a further condition on the tail function, the so-called third order condition. Two specific examples of kernel statistics are studied in greater depth, and their asymptotic behavior illustrated numerically. The finite sample properties are examined by means of a simulation study.

1. Introduction

A distribution is said to be of Pareto-type if for some $\gamma > 0$ its survival function is of the form

$$1-F(x) = x^{-1/\gamma}\ell_F(x), \quad x > 0, \tag{1}$$

where ℓ_F denotes a slowly varying function at infinity, i.e.

$$\frac{\ell_F(\lambda x)}{\ell_F(x)} \to 1 \quad \text{as } x \to \infty \text{ for all } \lambda > 0. \tag{2}$$

The parameter γ, called the extreme value index, clearly governs the first order tail behavior, with larger values of γ indicating heavier tails. The Pareto-type model can also be stated in an equivalent way in terms of the tail quantile function U, where $U(x) := \inf\{y : F(y) \geq 1-1/x\}$, $x > 1$, as follows:

$$U(x) = x^{\gamma}\ell_U(x), \tag{3}$$

with ℓ_U again a slowly varying function at infinity (Gnedenko, 1943).

* Corresponding author.
E-mail addresses: yuri.goegebeur@stat.sdu.dk (Y. Goegebeur), jan.beirlant@wis.kuleuven.be (J. Beirlant), tdewet@sun.ac.za (T. de Wet).
[1] This author's research was supported by a grant from the Danish Natural Science Research Council.
[2] This author's research was partially supported by the National Research Foundation.

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2010.03.029

Pareto-type models have important practical applications, and are used rather systematically in certain branches of non-life insurance, as well as in finance (stock returns), telecommunications (file sizes, waiting times), geology (diamond values, earthquakes), and many others. In the analysis of Pareto-type models, the estimation of the tail parameter γ, and the subsequent estimation of extreme quantiles assume a central position. Several estimators have been proposed in the literature, and their asymptotic distributions established, usually under a second order condition on the tail behavior (we refer to Beirlant et al., 2004; de Haan and Ferreira, 2006, for recent overviews of such estimation procedures).

Second order condition ($\mathcal{R}$). *There exists a positive real parameter γ, a negative real parameter ρ and a function b with $b(t)\to 0$ for $t\to\infty$, of constant sign for large values of t, such that*

$$\lim_{t\to\infty}\frac{\ln U(tx)-\ln U(t)-\gamma\ln x}{b(t)}=\frac{x^{\rho}-1}{\rho},\quad \forall x>0.$$

Let us remark here that condition ($\mathcal{R}$) implies that $|b|$ is regularly varying of index ρ (Geluk and de Haan, 1987), and hence the parameter ρ determines the rate of convergence of $\ln U(tx)-\ln U(t)$ to its limit, being $\gamma\ln x$, as $t\to\infty$. Condition ($\mathcal{R}$) is not too restrictive; for instance, the important Hall class of Pareto-type models (Hall and Welsh, 1985) for which the tail quantile function is of the form

$$U(x)=Cx^{\gamma}(1+Dx^{\rho}+o(x^{\rho}))\quad (x\to\infty)$$

with $C>0$, $D\neq 0$ satisfies the second order condition with $b(x)=\rho Dx^{\rho}$. This class is quite broad and contains distributions like the Fréchet, Burr, Generalized Pareto (GP) and $|t|$, to name but a few.

In this paper we deal with the estimation of the second order parameter ρ in condition ($\mathcal{R}$). Adequate estimation of this parameter has practical relevance for at least two reasons. Firstly, the parameter ρ is of crucial importance for procedures developed to adaptively select the optimal number of upper order statistics in tail index estimation, as can be seen in the papers by Hall and Welsh (1985), Beirlant et al. (1996), Drees and Kaufmann (1998), Guillou and Hall (2001), among others. Secondly, recent research on tail index estimation focuses on the development of improved estimators for γ by explicitly estimating the dominant term of the asymptotic bias. As this bias term depends on ρ, an a priori estimation of this parameter is needed in order to obtain the bias reduction. We refer to Beirlant et al. (1999), Feuerverger and Hall (1999), Beirlant et al. (2002), Gomes and Martins (2002), and more recently to Gomes et al. (2008), for examples of such bias-corrected estimators.

Under the second order framework, Gomes et al. (2002) and Fraga Alves et al. (2003) consider the generalized Hill statistic of Dekkers et al. (1989), given by

$$M_n^{(\alpha)}(k):=\frac{1}{k}\sum_{j=1}^{k}[\ln X_{n-j+1,n}-\ln X_{n-k,n}]^{\alpha},$$

where α is some positive tuning parameter and $X_{1,n}\leq\cdots\leq X_{n,n}$ denote the order statistics of a random sample from a distribution function satisfying ($\mathcal{R}$), and construct various ratios of linear combinations of powers of such statistics, which, under appropriate conditions on k and n, yield consistent estimators for functions of ρ. Inversion of the latter yields then ultimately consistent ρ estimators. Unlike earlier papers on ρ estimation, these authors carry the analytical derivation one step further and establish, besides consistency, also asymptotic normality, and this under a third order condition on the tail behavior. The Fraga Alves et al. (2003) estimator is to date generally considered to be the best working one in practice, and will be a reference method in our numerical and simulation experiments.

The approach to ρ estimation followed in this paper is based on the kernel goodness-of-fit statistics for Pareto-type behavior introduced in Goegebeur et al. (2008). In fact the latter paper proposed already a simple estimator for ρ, although the asymptotic behavior of it was not explored there. Inspired by the basic idea of Gomes et al. (2002) and Fraga Alves et al. (2003), we take a ratio of a ratio of a difference of two power transformed kernel goodness-of-fit statistics to obtain a quantity for which the dominant term of the asymptotic expansion is a function of ρ, and obtain the estimator by inversion of the latter. The approaches of Gomes et al. (2002) and Fraga Alves et al. (2003) were recently extended by Ciuperca and Mercadier (2010) by the inclusion of a weight function. The main difference with the above mentioned papers is that our approach to ρ estimation is based on the scaled log-spacings of successive order statistics introduced in Beirlant et al. (1999) rather than the ordinary excesses over a high threshold. Thus, we introduce a general and flexible class of estimators for ρ, for which the consistency and asymptotic normality can be derived under some rather mild conditions on the kernel functions.

The remainder of this paper is organized as follows. In the next section we develop the estimators and study their asymptotic properties, namely consistency and asymptotic normality. In this, we pay special attention to two particular examples of kernel functions, namely the power and log kernel functions. Numerical and small sample simulation results are reported in Section 3. The proofs of the results are deferred to the Appendix.